

METHOD OF MOTION ESTIMATION FOR TRANSMISSION
COST REDUCTION OF MOTION VECTORS

The invention relates to a method of movement
5 estimation applied to MPEG-type video coding.

The majority of movement-estimation algorithms
implemented in video coding use the technique of "block
matching".

The image is segmented into blocks of size $N \times N$,
10 called macroblocks, and the estimator searches for the
vector minimizing the difference between a block of the
current image and a block of the reference image. This
difference is generally an MSE (Mean Square Difference)
or MAE (Mean Absolute Difference) calculated on the
15 luminance pixels.

This type of estimator can supply a heteroge-
neous movement field since it is based on the varia-
tions of luminance and not on the actual movement in
the sequence. This may entail an overhead for the cod-
20 ing of the vectors by the coder, the coding generally
being of differential type, and thus a reduction in
performance.

The object of the invention is to remedy the
abovementioned drawbacks.

25 Its subject is a method of movement estimation
including segmentation of the video image into image
blocks, movement estimation per image block in order to
obtain a movement vector field, characterized in that
it includes a stage of reassignment of a vector to a
30 block by selecting one movement vector from among N
predominant vectors belonging to the vector field.

According to one particular implementation, for
a predominant vector, second-order regional maxima are
detected so as not to be taken into account during the
35 selection of the other predominant vectors.

According to another implementation, the pre-
dominant vectors are selected in each of the four
directions.

According to a particular implementation of the method, the selection of the reassigned vector is based on the value of the inter-displaced-image difference (DFD).

5 A particular characteristic of the invention consists in adopting the zero vector if the DFDs associated with the N predominant vectors are greater than the DFD associated with the original vector, or in actually keeping the original vector if the DFDs associated with the N predominant vectors are greater than
10 the weighted DFD associated with the original vector.

 According to another implementation of the method, the selection of the reassigned vector is based on the calculation of the activity (spatial gradient)
15 in the inter-image difference block (current block - estimated block). If the activities corresponding to the N predominant vectors are greater than the activity corresponding to the original vector, the zero vector is adopted. If the activities corresponding to the N
20 predominant vectors are greater than the weighted activity corresponding to the original vector, the original vector is kept.

 According to another particular implementation of the method, for each image, the predominant vectors
25 are chosen from among the field of vectors of the current image and the field of vectors of at least one preceding image.

 By virtue of the invention, the movement vector fields calculated by an estimator of the "block matching"
30 type can be homogenized.

 The characteristics and advantages of the invention will emerge better from the following description, given by way of example and by reference to the attached figures, in which:

35 - Figure 1 represents a histogram of the movement vectors,

 - Figure 2 represents a regional-maxima search window,

- Figure 3 represents an example of median filtering,

- Figure 4 represents an example of the preceding image vectors being taken into account,

5 - Figure 5 represents movement-vector fields during a zoom,

- Figure 6 represents various types of movement which may be detected.

10 The homogenization of the vector field is obtained via a method of conditional reassignment.

The vectors, associated with the images of a sequence, are calculated and stored by the estimator.

15 In order to carry out processing on the vectors, a two-dimensional histogram is constructed with dimensions of 512*512 in which the coordinates represent the values (dx, dy) which are the values of the horizontal and vertical components of these vectors.

20 Figure 1 represents, on the left-hand part, an image consisting of macroblocks to which the movement vectors are allocated and, on the right-hand part, the corresponding histogram.

Choice of predominant vectors

25 In order to make the movement field more homogeneous, the idea is to adopt a certain number of vectors, which is fixed in the first place by the user. This number will be larger in proportion to the heterogeneity of the movements.

30 The first solution consists in adopting the N vectors corresponding to the highest frequencies of appearance.

35 Another possibility is to stipulate that the algorithm choose N/4 predominant vectors in each of the four orientation planes. This solution can be adopted as an option, as an output criterion upon detection of zoom in the sequence. This is because such a phenomenon entails distribution in all directions of the vector field.

The last solution envisaged is to carry out detection of the regional maxima. This is because the

problem, in the first solution, is that it is possible to have several contiguous maxima, which do not confer enormous advantages compared with the fact of adopting fewer of them.

5 The histogram is therefore scanned, rejecting those vectors among the N predominant vectors appearing in the vicinity of other more predominant vectors. Thus the existence of these second-order maxima is identified by looking at the histogram to see whether two
10 maxima lie in the same window, for example with dimensions 3*3.

 Figure 2 represents such a window, referenced 1, for searching for regional maxima, this window being centred around the predominant vector adopted (dX, dY),
15 the number of occurrences of which is n.

Choice of the vector allocated to a macroblock MB. Re-assignment

- Method of the DFD

 Once the predominant vectors have been
20 extracted, a criterion remains to be found for reassigning each of these vectors to each MB. Since the movement estimator uses the criterion of the minimum DFD (Displaced-Frame Difference) to calculate the movement vectors, it seems useful to use this criterion to
25 find the best possible correspondence between the vectors adopted and the macroblocks of the image to be processed.

 After ordering the vectors in increasing order of their frequency of appearance, the calculation of
30 DFD associated with each of these vectors is carried out for each MB. This calculation can be expressed simply by the following formula:

$$Dfd(i, j) = \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} |MBCurrent(i+k, j+l) - MBReference(i+k+dy, j+l+dx)|$$

35 in which (i, j) are the coordinates of the MB to be processed;

N (= 16) is the size of the MB;

(dx, dy) are the components of the vector to be tested, belonging to [-128; +127.5].

It is important, before applying this formula, to check that the vector to be tested does not point outside the reference image. If no vector is suitable, then the zero vector is assigned.

Hence the vector corresponding to the minimum DFD is assigned to each MB.

- Gradient method

This consists in seeking, for each MB of the "difference" image consisting of the predicted reference image and of the current image, the vector corresponding to the minimum gradient which gives information on the local activity of the MB (of horizontal and vertical gradient type).

$$\text{MB_gradient} = \sum_{\substack{4\text{luma} \\ \text{blocks}}} \text{block_active}$$

with:

$$\text{block_active} = \text{MAX} \left(\text{MAX}_{i,j=0}^{i=6,j=7} |x(i,j) - x(i+1,j)|, \text{MAX}_{i,j=0}^{i=7,j=6} |x(i,j) - x(i,j+1)| \right)$$

Enhancement of the reassignment

DFD/Gradient criterion

In order to keep certain movements, relating to objects of small size, the following criterion is defined:

If, after application of the DFD method, the vector adopted for an MB generates a DFD greater than the weighted original DFD, the original vector is kept.

Likewise, regarding the method of the gradient, for each MB obtained after inter-image difference, the gradient obtained by reassignment is compared with the gradient of the original vector. If the weighted original gradient is less than the new gradient, the original vector is kept.

Filtering applied to the movement vectors

In order to make the vector fields more homogeneous, other criteria may be used, namely spatial or temporal filtering.

5 - Spatial filtering

The filter adopted is the two-dimensional 3*3 median filter:

the principle is explained below in the light of Figure 3 which represents an image referenced 2 before filtering and an image referenced 3 after filtering. The vector referenced 4 is the vector to be processed.

The vertical and horizontal neighbours of the components of the MB in question are ordered along each direction (dx, dy), then the median value of each component is taken. Next the various DFDs associated with each MB are compared, in the case in which either one component is filtered, or both, or no component is filtered. Hence the vector corresponding to the minimum DFD is chosen, the original DFD, obviously, being weighted.

- Temporal filtering

The idea of temporal coherence is to take account, in the reassignment of the vectors of an image, of the movement fields of the preceding images; this is done with a view to limiting the disparity in the movements from one image to another.

To begin with, we will detail the principle of temporal filtering of Forward vectors (deferred-movement vectors).

Spatio-temporal histogram of Forward vectors:

In order to take account of the various histograms, scaling of the vectors is carried out at a first stage, then weighting of the occurrences which is a function of the position of the various histograms with respect to the histogram processed.

Hence, for the P image of Figure 4, it is possible to add to the histogram of original vectors, the occurrences of which have been weighted by a factor 3,

the occurrences of the vectors of the first B (the amplitude of which has been multiplied by 3) which are weighted by a factor 1 as well as the occurrences of the vectors of the second B (the amplitude of which has been multiplied by 3/2) which are weighted by a factor 2.

Temporal coherence should be relevant when uniform movements are present, and breaks in movement (change of scene) are not present.

Case of Backward vectors (anticipated-movement vectors)

It would be logical to think that, if there are uniform "Forward" movements from one image to the next, they would also be present in the case of the "Backward" vectors associated with the B images. In order to filter the latter, it must not be forgotten that the Backward vectors are based on the P or the I which will follow the B in question. Hence, for the first B, it may be thought that its Backward vectors will be twice as large as the Backward vectors associated with the second B. Scaling is carried out on the vectors of the latter by a factor of 2, and the weighted occurrences will be added, in the histogram associated with the first B.

Detection of uniform field

The idea of applying the reassignment with N vectors on sequences with multidirectional movements such as a zoom, for example, is not relevant. This is because, in this fairly specific case, the fact of adopting only N predominant vectors does not make it possible conveniently to process the fields consisting of multiple vectors.

Figure 5 represents the image of the vectors during the zoom. It can easily be seen that the disparity in the field does not allow any such uniformity.

It is therefore decided to detect, in the first place, a field in which the vectors are uniformly distributed, either unilaterally, or in all directions (zoom). This detection is conveyed by a standard devia-

tion of the first predominant vector close to the average standard deviation calculated from among the N predominant vectors. This is expressed as:

if $\sigma_1 \leq \text{threshold} * \sigma_{\text{average}} \Rightarrow \text{uniform field present}$
5 in which the threshold is fixed by the user (threshold = 1.34 for example).

Examples relating to the types of movements which are successfully detected are represented in Figures 6a, b, c, d.

10 The objective is, at present, not to apply the algorithm when cases (c) and (d) are present. These cases have still to be distinguished from cases (a) and (b). To do that the average values of the dx and dy movements are examined, from among the N adopted, and
15 it is seen whether they are close to zero. This is because it may be observed that the movements in a zoom seem to cancel out if they are added, in contrast to unilateral movement. A maximum difference of five pixels can be set for dx, dy.

20 Limitation on the temporal filtering

It is useful not to have to filter the histograms temporally in the event of breaks in movement. It is possible:

- to store the histogram of initial or reas-
25 signed vectors for a P-type image;

- at the next P-type image, P (t), the new "image" vectors are compared. If they differ too much from their counterparts arising from P (t - n), the original vectors are kept.

30 Choice of the Number of Predominant Vectors

The number of vectors necessary may be decided automatically and dynamically, in such a way that, for sequences with random movements (for example a sporting sequence), there are more vectors than for sequences
35 with uniform movements ("train").